



## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **11224256 A**

(43) Date of publication of application: 17 . 08 . 99

(51) Int. Cl. **G06F 17/30**(21) Application number: **10024748**(22) Date of filing: **05 . 02 . 98**(71) Applicant: **NIPPON TELEGR & TELEPH  
CORP <NTT>**(72) Inventor: **YOSHIDA KATSUHIKO  
YAMAMOTO SHUICHIRO  
TAKADA SHINICHI**(54) **INFORMATION RETRIEVING METHOD AND  
RECORD MEDIUM RECORDING INFORMATION  
RETRIEVING PROGRAM**

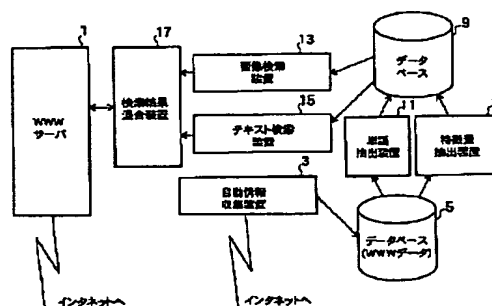
and key word and their points, and retrieval result is obtained according to the order.

COPYRIGHT: (C)1999,JPO

(57) Abstract:

**PROBLEM TO BE SOLVED:** To provide an information retrieving method which efficiently and accurately retrieves information containing an image and to provide a record medium where an information retrieving program is recorded.

**SOLUTION:** An automatic information collecting device 3 collects data including images on WWW's, a feature quantity extracting device 7 extracts feature quantities of the gathered data, and a word extracting device 11 extracts words in the vicinity of image information, imparts points depending upon how close to the image the respective words are, and stores the data and words in a database 9 so that they are made to correspond to the data feature quantities and points; when retrieval is performed, the feature quature quantity of a key image is extracted to obtain an approximation order from the feature quantity of data in the database 9, images relating to a key word are sequenced according to images in the database, the approximation order of the key image, and words in database matching the key image



Japanese Patent Application Laid-Open No. 224256/1999

(Tokukaihei 11-224256)

(54)[Title of the Invention] Information retrieving method and a recording medium having such an information retrieving program.

(57)[Abstract]

[Object] To provide an information retrieving method for retrieving information containing images efficiently and appropriately and a recording medium having such an information retrieving program.

[Means to solve the problems] A method having the steps of: collecting data containing an image on the WWW by using an automatic information collecting device 3; extracting the feature amount of this collected data by using a feature-amount extraction device 7; extracting words in the vicinity of the image information by using a word extraction device 11; adding points to the words in the order in which they are closer to the image so as to store them in a data base 9 with the data and words being associated with the data feature amount and the points; upon retrieval, extracting the feature amount of a key image, while finding the order of approximation from the feature amount of data within the data base 9; and with respect to the key image and a key word, based upon the order of approximation between the key image and the image within the data base as well as the key word and those words within the data base that are

coincident with the key word and their points, prioritizing the images relating to the key word, thereby obtaining the results of the retrieval based upon the order of priority.

[EMBODIMENTS]

[0013] FIG. 1 is a block diagram that shows a construction of an information retrieving system for executing an information retrieving method in accordance with one embodiment of the present invention. The information retrieving system shown in this Figure is provided with a WWW server 1 connected to the Internet so as to carry out normal server functions, and an automatic information collecting device 3 that is also connected to the Internet, and collects data including images on the WWW. The data including images collected by the automatic information collecting device 3 is temporarily stored in a data base 5.

[0014] The collected data, stored in the data base 5, is read out by a feature amount extraction device 7 so that the amount of feature is extracted, and stored in another data base 9. Moreover, the collected data, stored in the data base 5, is also read out by a word extraction device 11, and its HTML language of the WWW is analyzed so that words located in the vicinity of the image information are extracted, and based upon the assumption that the closer to the image information the extracted word is located, the more closely the word is related to the image, higher points are applied to those words in the order in which they are closer to the image, and each word with

the point is stored in the data base 9 in association with the corresponding image information and the amount of feature of the image. In other words, in the data base 9, with respect to each piece of image information, the amount of feature of the image, each word related to the image and the point of each word are stored in association with each other.

[0015] The data including the image information stored in the data base 9 is used when desired image information is retrieved based upon a key image by an image retrieving device 13 and when desired text information is retrieved based upon the key word by a text retrieving device 15. The image retrieving device 13 extracts the amount of feature of the key image inputted by the user of the Internet through the WWW server 1, and this extracted amount of feature of the key image is collated with the amount of feature of the image inside the data server 9 so that the order of approximation is outputted. Moreover, the text retrieving device 15 collates the key word inputted by the Internet user through the WWW server 1 with the words inside the data base 9, and based upon the coincided word and its points, prioritizes images related to the key word, and outputs information of the order of priority related to the image.

[0016] The information of the order of approximation of images related to the key image outputted from the image retrieving device 13 and the information of the order of priority of images related to the key word outputted from the text retrieving device 15 are supplied to a retrieval result mixing device 17, which outputs the results of retrieval based upon the

information of the order of approximation of images related to the key image and the information of the order of priority of images related to the key word. The results of retrieval are sent to the user from the WWW server 1.

[0017] Next, referring to a system structural drawing in FIG. 2 and a flow chart in FIG. 3, an explanation will be given of an information collecting process carried out in the information retrieving system arranged as described above.

[0018] As illustrated in FIG. 2, the information retrieving system of the present embodiment, shown in FIG. 1, is connected to a plurality of Web sites 23a, 23b, ..., 23n, through the Internet 21, so that texts written in HTML and image information, stored in their data bases 25a, 25b ...25n, are collected by the automatic information collecting device 3 through the Internet 21, and stored in the data base 5 as described above.

[0019] In this image collecting process, as shown in the flow chart of FIG. 3, first, the automatic information collecting device 3 of the information retrieving system specifies a desired Web site 23 from which information is to be collected, and information is collected from the data base 25 of the specified Web site 23, and the resulting information is stored in the data base 5 (step S11). After such an information collecting process has been carried out on all the Web sites (step S13), image files are selected from the collected information, and the feature amount extraction device 7 extracts the amount of feature of each of these images, and the resulting data is stored in a data base 9 (step S15). Then,

the key word is extracted by analyzing the HTML file, and a point is applied to the extracted key word depending on the distance from a tag specifying the image, and stored in the data base 9 (step S17).

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-224256

(43) 公開日 平成11年(1999) 8月17日

(51) Int.Cl.<sup>8</sup>

G 0 6 F 17/30

識別記号

F I

G 0 6 F 15/403

3 5 0 C

15/40

3 1 0 F

3 7 0 B

15/401

3 1 0 C

審査請求 未請求 請求項の数 2 O L (全 7 頁)

(21) 出願番号 特願平10-24748

(22) 出願日 平成10年(1998) 2月5日

特許法第30条第1項適用申請有り 1997年11月28日 社団法人電子情報通信学会主催の「電子情報通信学会技術研究報告会 知能ソフトウェア工学研究会」において文書をもって発表

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 吉田 勝彦

東京都新宿区西新宿三丁目19番2号 日本電信電話株式会社内

(72) 発明者 山本 修一郎

東京都新宿区西新宿三丁目19番2号 日本電信電話株式会社内

(72) 発明者 ▲高▼田 信一

東京都新宿区西新宿三丁目19番2号 日本電信電話株式会社内

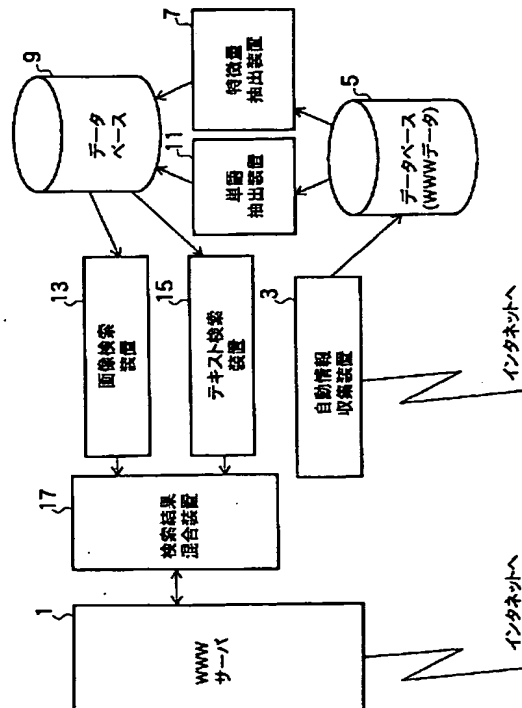
(74) 代理人 弁理士 三好 秀和 (外1名)

(54) 【発明の名称】 情報検索方法および情報検索プログラムを記録した記録媒体

(57) 【要約】

【課題】 画像を含んだ情報を効率的かつ適確に検索し得る情報検索方法および情報検索プログラムを記録した記録媒体を提供する。

【解決手段】 自動情報収集装置 3 で WWW 上の画像を含むデータを収集し、この収集データの特徴量を特徴量抽出装置 7 で抽出し、単語抽出装置 11 で画像情報の近傍の単語を抽出し、各単語に対して画像に近い順に高い点数を付与し、データと単語とをそれぞれのデータ特徴量と点数とに対応付けてデータベース 9 に格納し、検索時は、キー画像の特徴量を抽出し、データベース 9 内のデータの特徴量から近似順位を求め、キー画像とキーワードに対してデータベース内の画像とキー画像の近似順位およびキーワードとデータベース内の単語との一致したものとその点数からキーワードに関係する画像を順位付けし、該順位に基づいて検索結果を得る。



## 【特許請求の範囲】

【請求項 1】 WWW (World Wide Web) 上の画像を含む情報を検索する情報検索方法であって、  
予め WWW 上の画像を含むデータを収集し、この収集したデータの特徴量を抽出し、  
前記データ中の画像情報の近傍にある単語を抽出し、該単語のそれぞれに対して画像に近い順に高い点数を付与し、  
前記データと単語とをそれぞれのデータ特徴量と点数とに対応付けてデータベースとして格納し、  
検索時は、キー画像の特徴量を抽出し、前記データベース内のデータの特徴量から近似順位を求め、キー画像とキーワードに対して前記データベース内の画像とキー画像の近似順位およびキーワードとデータベース内の単語との一致したものとその点数から前記キーワードに関係する画像の順位付けを行い、該順位に基づいて検索結果を得ることを特徴とする情報検索方法。

【請求項 2】 WWW (World Wide Web) 上の画像を含む情報を検索する情報検索プログラムを記録した記録媒体であって、  
予め WWW 上の画像を含むデータを収集し、この収集したデータの特徴量を抽出し、  
前記データ中の画像情報の近傍にある単語を抽出し、該単語のそれぞれに対して画像に近い順に高い点数を付与し、  
前記データと単語とをそれぞれのデータ特徴量と点数とに対応付けてデータベースとして格納し、  
検索時は、キー画像の特徴量を抽出し、前記データベース内のデータの特徴量から近似順位を求め、キー画像とキーワードに対して前記データベース内の画像とキー画像の近似順位およびキーワードとデータベース内の単語との一致したものとその点数から前記キーワードに関係する画像の順位付けを行い、該順位に基づいて検索結果を得ることを特徴とする情報検索プログラムを記録した記録媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、WWW (World Wide Web) 上の画像を含む情報を検索する情報検索方法および情報検索プログラムを記録した記録媒体に関する。

## 【0002】

【従来の技術】 インターネット上には WWW の Web サイトと呼ばれる装置を設けることにより情報を公開することができる。この Web サイトは世界中に多数存在し、HTML (Hyper Text Markup Language) と呼ばれるテキスト情報で記述されており、リンクを辿ることにより情報を芋蔓式に見ていくことができる。しかしながら、このままでは情報の検索性が悪いので、テキスト情報を収集し、キーワードを利用してテキスト情報を検索するサービスが実現されている。

【0003】 また、HTML ではテキスト以外の情報についても参照関係を記述することによりマルチメディアを利用した文章を記述することができる。現在、このようなテキスト以外の情報を検索する場合、関係するキーワードを考え、このキーワードを用いて検索した結果のコメント欄の内容や、実際にその情報を見てから判断している。

【0004】 更に、画像検索技術には画像同士の色や形の近似性から類似画像を検索する方法や、画像に人手でキーワードを付与し、このキーワードを基に検索する方法があるが、WWW のテキストや画像を含めた形で情報検索に利用されている例は無い。

## 【0005】

【発明が解決しようとする課題】 WWW に登録されている情報は、マルチメディアで構成されているのがほとんどであり、このようなテキスト以外の情報も含めて検索したい場合には、キーワード検索の結果の内容を一通り見て判断する必要がある。この場合、元になるイメージ画像を持っていた場合でも、その画像を活用することはできない。例えば、蝶の画像を持っており、その蝶の種類、生息地等を知りたい場合、蝶であるということがわかっていれば、蝶と言うキーワードから検索を始め、画像が登録されているページを検索結果から手作業で探していき、手元の写真と登録されている画像を見比べて同じと思われるものを見つけ、それから必要な説明が書かれているか確認すると言った複雑な作業を行う必要があるという問題がある。

【0006】 このような検索結果は場合によっては数千件検索されることがある。また、蝶であることすら知らなければ検索することはより困難になるという問題がある。

【0007】 本発明は、上記に鑑みてなされたもので、その目的とするところは、画像を含んだ情報を効率的かつ適確に検索し得る情報検索方法および情報検索プログラムを記録した記録媒体を提供することにある。

## 【0008】

【課題を解決するための手段】 上記目的を達成するため、請求項 1 記載の本発明は、WWW (World Wide Web) 上の画像を含む情報を検索する情報検索方法であって、予め WWW 上の画像を含むデータを収集し、この収集したデータの特徴量を抽出し、前記データ中の画像情報の近傍にある単語を抽出し、該単語のそれぞれに対して画像に近い順に高い点数を付与し、前記データと単語とをそれぞれのデータ特徴量と点数とに対応付けてデータベースとして格納し、検索時は、キー画像の特徴量を抽出し、前記データベース内のデータの特徴量から近似順位を求め、キー画像とキーワードに対して前記データベース内の画像とキー画像の近似順位およびキーワードとデータベース内の単語との一致したものとその点数から前記キーワードに関係する画像の順位付けを行い、該



順位に基づいて検索結果を得ることを要旨とする。

【0009】請求項1記載の本発明にあっては、WWW上から収集した画像を含むデータの特徴量を抽出し、画像情報の近傍の単語を抽出し、各単語に対して画像に近い順に高い点数を付与し、データと単語とをそれぞれのデータ特徴量と点数とに対応付けてデータベースとして格納し、検索時は、キー画像の特徴量を抽出し、データベース内のデータの特徴量から近似順位を求め、キー画像とキーワードに対してデータベース内の画像とキーワードの近似順位およびキーワードとデータベース内の単語との一致したものとその点数からキーワードに関係する画像を順位付けし、該順位に基づいて検索結果を得るため、画像からの画像情報の検索が可能となるとともに、キーワードからの画像情報検索、画像からの音声情報の検索、テキストからの画像情報の検索、画像とテキストを同時に利用した情報検索等のように柔軟な検索が可能となる。

【0010】また、請求項2記載の本発明は、WWW（World Wide Web）上の画像を含む情報を検索する情報検索プログラムを記録した記録媒体であって、予めWWW上の画像を含むデータを収集し、この収集したデータの特徴量を抽出し、前記データ中の画像情報の近傍にある単語を抽出し、該単語のそれぞれに対して画像に近い順に高い点数を付与し、前記データと単語とをそれぞれのデータ特徴量と点数とに対応付けてデータベースとして格納し、検索時は、キー画像の特徴量を抽出し、前記データベース内のデータの特徴量から近似順位を求め、キー画像とキーワードに対して前記データベース内の画像とキーワードの近似順位およびキーワードとデータベース内の単語との一致したものとその点数から前記キーワードに関係する画像の順位付けを行い、該順位に基づいて検索結果を得ることを要旨とする。

【0011】請求項2記載の本発明にあっては、WWW上から収集した画像を含むデータの特徴量を抽出し、画像情報の近傍の単語を抽出し、各単語に対して画像に近い順に高い点数を付与し、データと単語とをそれぞれのデータ特徴量と点数とに対応付けてデータベースとして格納し、検索時は、キー画像の特徴量を抽出し、データベース内のデータの特徴量から近似順位を求め、キー画像とキーワードに対してデータベース内の画像とキーワードの近似順位およびキーワードとデータベース内の単語との一致したものとその点数からキーワードに関係する画像を順位付けし、該順位に基づいて検索結果を得る情報検索プログラムを記録媒体として記録しているため、該記録媒体を利用して、その流通性を高めることができる。

【0012】

【発明の実施の形態】以下、図面を用いて本発明の実施の形態について説明する。

【0013】図1は、本発明の一実施形態に係る情報検

索方法を実施する情報検索システムの構成を示すブロック図である。同図に示す情報検索システムは、インターネットに接続され、通常のサーバ機能を行うWWWサーバ1および同様にインターネットに接続され、WWW上の画像を含むデータを収集する自動情報収集装置3を有する。この自動情報収集装置3で収集された画像を含むデータは一旦データベース5に格納される。

【0014】このデータベース5に格納された収集データは、特徴量抽出装置7によって読み出されて、その特徴量が抽出され、別のデータベース9に格納される。また、データベース5に格納された収集データは、同様に単語抽出装置11によって読み出され、WWWのHTML文章を解析されて、画像情報の近傍にある単語を抽出し、この抽出した単語が画像情報の近傍に記述されている程、その画像に関係の深い単語であると仮定し、この単語の各々に対して画像に近い順に高い点数を付与し、各単語をその点数とともに該当する画像情報および該画像の特徴量に対応付けてデータベース9に格納する。すなわち、データベース9は、各画像情報に関連して該画像の特徴量、該画像に関係する各単語および各単語の点数をそれぞれ対応させて格納している。

【0015】データベース9に格納された画像情報を含むデータは、画像検索装置13でキー画像に基づいた所望の画像情報が検索されるとともに、テキスト検索装置15によってキーワードに基づいた所望のテキスト情報が検索される。画像検索装置13は、WWWサーバ1を介してインターネット利用者から入力されるキー画像の特徴量を抽出し、この抽出したキー画像の特徴量をデータベース9内の画像の特徴量と照合して、近似順位を出力する。また、テキスト検索装置15は、WWWサーバ1を介してインターネット利用者から入力されるキーワードをデータベース9内の単語と照合して一致した単語とその点数から該キーワードに関係する画像の順位付けを行い、この画像の順位付け情報を出力する。

【0016】画像検索装置13から出力されるキー画像に対する画像の近似順位情報およびテキスト検索装置15から出力されるキーワードに関係する画像の順位付け情報は検索結果混合装置17に供給され、ここでキー画像に対する画像の近似順位情報とキーワードに関係する画像の順位付け情報に基づいて検索結果を出力する。この検索結果はWWWサーバ1から利用者に送出される。

【0017】次に、以上のように構成される情報検索システムにおける情報収集処理について図2に示すシステム構成図および図3に示すフローチャートを参照して説明する。

【0018】図1に示した本実施形態の情報検索システムは、図2に示すように、インターネット21を介して複数のWebサイト23a、23b、・・・、23nに接続され、これらの各Webサイト23a、23b、・・・、23nにそれぞれ設けられているデータベース2

5 a, 2 5 b, . . . , 2 5 n に格納されている HTML で記述されたテキストおよび画像情報をインターネット 2 1 を介して自動情報収集装置 3 で収集し、上述したようにデータベース 5 に格納するようになっている。

【0019】この情報収集処理においては、図 3 にフローチャートで示すように、まず情報検索システムの自動情報収集装置 3 が情報を収集したい所望の Web サイト 2 3 を指定し、この指定した Web サイト 2 3 のデータベース 2 5 から情報を収集し、ファイルとしてデータベース 5 に格納する（ステップ S 1 1）。このような情報収集をすべての Web サイトについて行うと（ステップ S 1 3）、この収集した情報から画像ファイルを選択し、この画像の特徴量を特徴量抽出装置 7 で抽出し、データベース 9 に格納する（ステップ S 1 5）。それから、HTML ファイルを解析してキーワードを抽出し、この抽出したキーワードに画像を指定しているタグからの距離によって得点を付与し、データベース 9 に格納する（ステップ S 1 7）。

【0020】更に具体的に、図 4 に示す HTML 文章例を参照しながら、図 1 に示す実施形態の作用について説明する。

【0021】情報検索システムの自動情報収集装置 3 はインターネット 2 1 を介して Web サイト 2 3 a, 2 3 b, . . . , 2 3 n の各々に順次アクセスし、その Web サイトのデータベース 2 5 a, 2 5 b, . . . , 2 5 n から WWW 上のデータを収集し、データベース 5 に格納する。それから、データベース 5 に格納された収集データのうち画像データの特徴量が特徴量抽出装置 7 によって抽出され、データベース 9 に格納される。また、HTML のテキストデータからは HTML のタグを解析し、画像ファイルを指定するタグ、" <img src = 画像ファイル名 " で始まるタグ、" <a href = 画像ファイル名 " で始まるタグの回りのテキストから単語を抽出する。図 4 に示す HTML 文章例では、四角で囲んだ所が画像タグであり、kodaitizu.jpg, hyoukeidai.wav, hyoukei.jpg は画像ファイルのファイル名である。

【0022】このようにタグの回りのテキストから単語を抽出する場合には、タグに近い距離に存在する単語に高い点数を付与して、データベース 9 に格納する。なお、点数の付け方は前もって決めておく。図 4 に示す HTML 文章例において、抽出された単語が、古代、文化、古代地図、紀元前、文明、文字板、楔形文字、板、文字、絵文字、楔形文字、紀元前であったとすると、画像 kodaitizu.jpg には、古代が 4 点、文化が 5 点、古代地図が 1 0 点、紀元前が 5 点、文明が 4 点、文字板が 3 点、音声 hyoukeidai.wav および画像 hyoukei.jpg にはそれぞれ、古代が 1 点、文化が 1 点、古代地図が 2 点、紀元前が 3 点、文明が 4 点、文字板が 5 点、楔形文字が 5 点、板が 4 点、文字が 3 点、絵文字が 2 点、楔形文字が 1 点、紀元前が 2 点となる。ここでは、画像タグの隣

の単語を 5 点、それ以降 1 点ずつ引いていき、1 点以降はすべて 1 点としている。ただし、画像 kodaitizu.jpg のタグの有効範囲は古代地図と言う単語も含んでいるので 1 0 点としている。なお、このように得られた結果が図 5 に示すように検索情報テーブルとしてデータベース 9 に格納されている。

【0023】以上のようにして、インターネットを介して各 Web サイトから画像情報を含むデータを収集した情報検索システムに対して、利用者が検索を行う場合において、画像のみをキーとして利用する場合には、利用者がキーにする画像をインターネットを介して情報検索システムの WWW サーバ 1 に送信する。WWW サーバ 1 は、この送信されてきたキー画像の特徴量を計算し、この計算した特徴量を既にデータベース 9 に格納されている画像の特徴量との近似関係を計算して、近い順に検索結果を表示する。そして、この検索結果からリンクを辿ることにより元の情報を見ることができる。

【0024】また、検索キーとしてキーワードを与えた場合には、データベース 9 に登録されているキーワードと情報登録時に付与した点数を基に計算を行い、点数の高い順に検索結果をキーワードと関連のある画像とともに表示する。そして、この検索結果からリンクを辿ることにより元の情報を見ることができる。

【0025】更に、画像とキーワードとが同時にキーとして与えられた場合には、画像に関しては上述した画像の検索と同様に近似関係を計算し、キーワードに関してもそれぞれ点数を計算する。そして、検索結果は画像が似ていて、キーワードの点数が高い順に表示する。この結果、テキストと画像を含めた検索および絞り込みが可能となる。

【0026】また、画像から音声を検索する場合は、画像を検索し、その結果の画像に付随する単語を基にその単語をキーワードとして検索し、音声のみを結果として返送する。更に、検索結果をキーとして選択し、再び検索を行うことができる。

【0027】

【発明の効果】以上説明したように、本発明によれば、WWW 上から収集した画像を含むデータの特徴量を抽出し、画像情報の近傍の単語を抽出し、各単語に対して画像に近い順に高い点数を付与し、データと単語とをそれぞれのデータ特徴量と点数とに対応付けてデータベースとして格納し、検索時は、キー画像の特徴量を抽出し、データベース内のデータの特徴量から近似順位を求め、キー画像とキーワードに対してデータベース内の画像とキー画像の近似順位およびキーワードとデータベース内の単語との一致したものとその点数からキーワードに係る画像を順位付けし、該順位に基づいて検索結果を得るので、例えばイメージ図から画像を検索し、その中の絵から必要とする画像を探し出したり、それに付随するテキスト情報を得ることができるというように画像が

らの画像情報の検索が可能となるとともに、更にキーワードからの画像情報検索、画像からの音声情報の検索、テキストからの画像情報の検索、画像とテキストを同時に利用した情報検索等のように柔軟な検索が可能となる。特に画像検索においては、キー画像とキーワードを同時に利用して検索を行うことにより絞り込んだ検索を行うことができる。

【0028】また、WWWの情報に音声や動画が含まれている場合には、画像と同様に動画、音声に対応する単語を対応付けておくことにより、画像を利用して音声を検索するといった別のメディア間での検索が可能となる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る情報検索方法を実施する情報検索システムの構成を示すブロック図である。

【図2】図1に示す情報検索システムにおける情報収集

処理を行う場合のシステム構成を示す図である。

【図3】図1に示す情報検索システムにおける情報収集処理を示すフローチャートである。

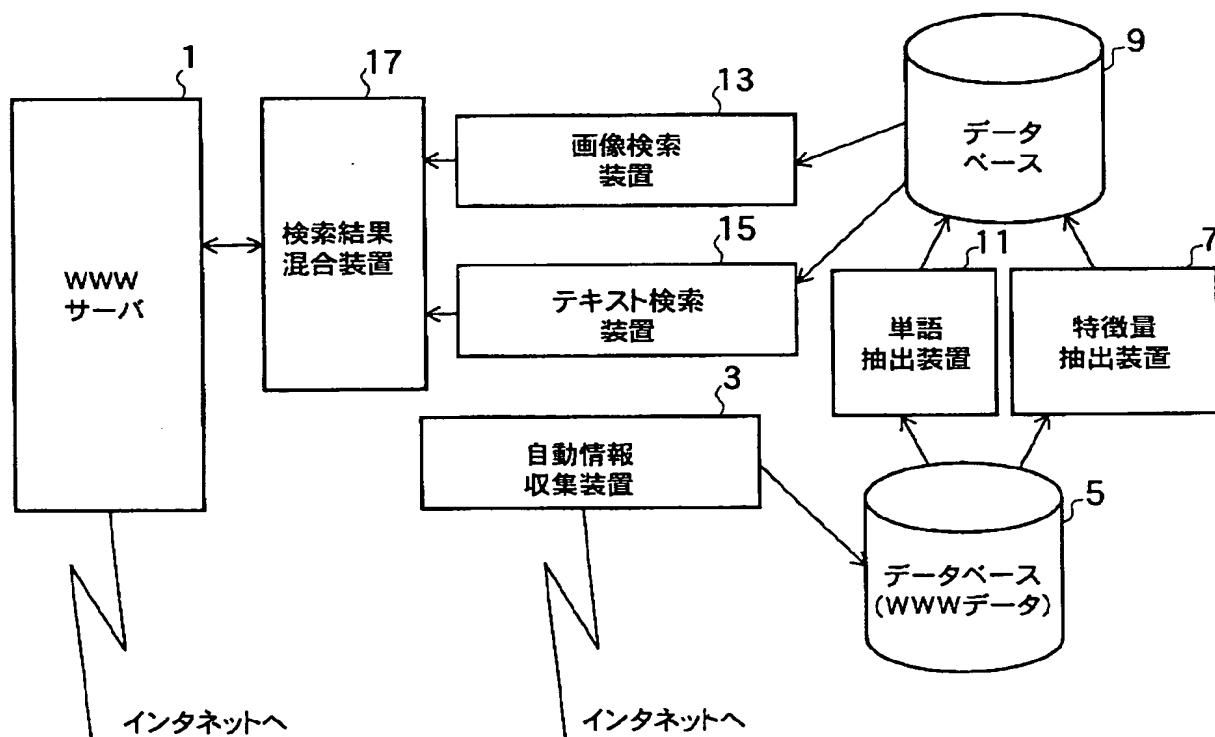
【図4】HTML文章例を示す図である。

【図5】図1に示す情報検索システムで収集した検索情報を格納するデータベースにおける検索情報テーブルの例を示す図である。

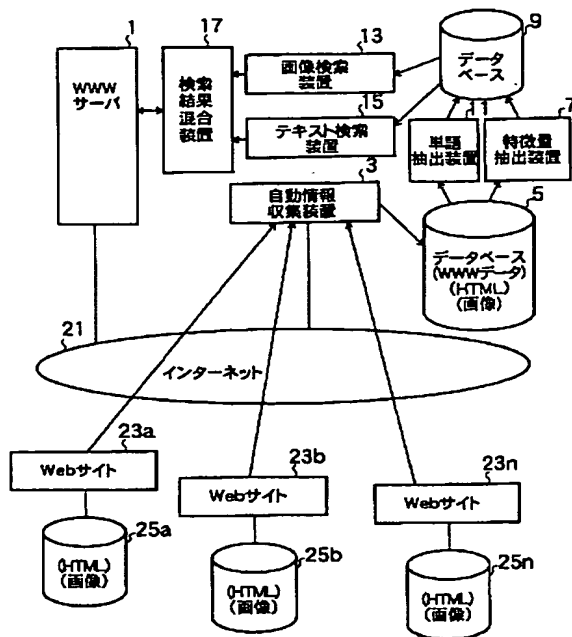
【符号の説明】

- 1 WWWサーバ
- 3 自動情報収集装置
- 5, 9 データベース
- 7 特徴量抽出装置
- 11 単語抽出装置
- 13 画像検索装置
- 15 テキスト検索装置
- 17 検索結果混合装置

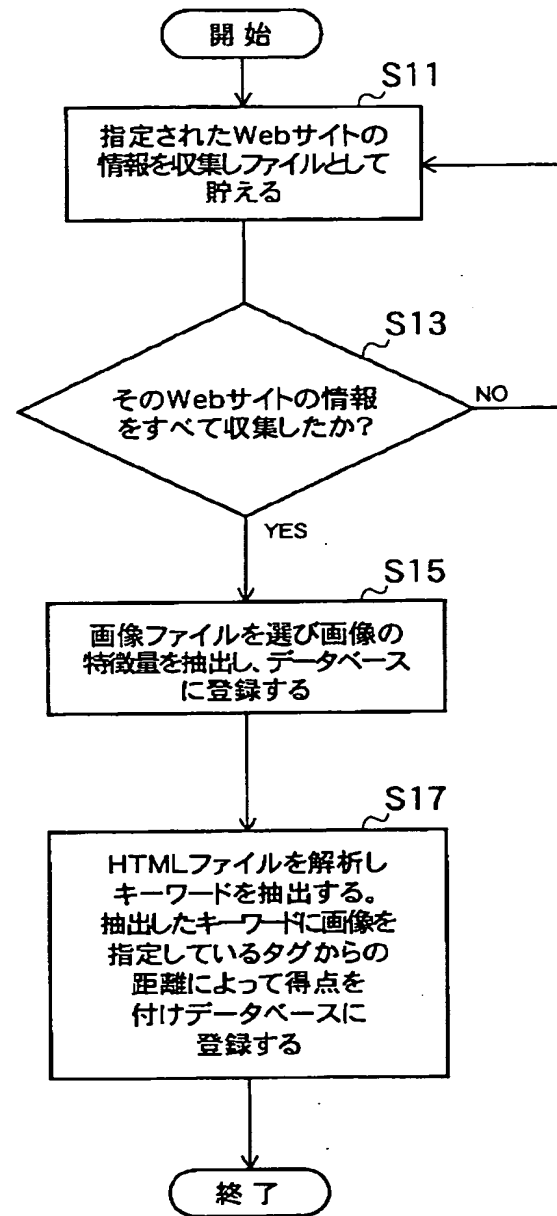
【図1】



【図 2】



【図 3】



【図 4】

## HTML文章例

```

<h2>古代の文化</h2>
<P><A HREF = "kodaitizu.jpg">古代地図</A></P>
<P>これは、紀元前10000年の文明のものでいわれている。そこで発見されたのがこの文
字板である。</P>
<P><A HREF = "hyoukeidai.wav"><IMG SRC = "hyoukei.jpg"></A><B>楔形文字の板
</B><BR>
<BR CLEAR = "ALL"><BR>
しかし、このころ使われていた文字は絵文字であり、楔形文字が使われるようになったの
は紀元前3000年頃であり、謎とされている。<BR>
<BR>

```

【図 5】

## 検索情報テーブル

種別	画像位置情報	単語	画像特徴量	情報位置 リンク情報
画像	http://aaa.aaa.aaa/ bbb/kodaitizu.jpg	古代 4、文化 5、古代 地図 10、紀元前 5、 文明 4、文字板 3	xxxxxxx x	http://aaa .aaa.aaa/ bbb/ddd .html
音声	http://aaa.aaa.aaa/ bbb/hyoukeidai.wav	古代 1、文化 1、古代 地図 2、紀元前 3、 文明 4、文字板 5、 楔形文字 5、板 4、 文字 8、絵文字 2、 楔形文字 1、紀元前 2		http://aaa .aaa.aaa/ bbb/ddd .html
画像	http://aaa.aaa.aaa/ bbb/hyoukei.jpg	古代 1、文化 1、古代 地図 2、紀元前 3、 文明 4、文字板 5、 楔形文字 5、板 4、 文字 3、絵文字 2、 楔形文字 1、紀元前 2	zzzzzzz zz	http://aaa .aaa.aaa/ bbb/ddd .html